Proceedings of ERAD 2006

Raindrop size distribution estimation from S-band polarimetric radar using a regularized neural network

Gianfranco Vulpiani[1], Frank S. Marzano[2], V. Chandrasekar[3], Alexis Berne[4], Remko Uijlenhoet[5]

[1] Meteo France, Trappes (France).
[2] University La Sapienza, Rome (Italy).
[3] Colorado State University, Fort Collins (Colorado) .
[4] LTHE, Grenoble (France).
[5] Wageningen University, Wageningen (The Netherlands).

1 Introduction

Radar rainrate estimates are prone to a high degree of uncertainty due to several error sources, among which the space-time variability of the raindrop size distribution (RSD). Polarimetric radar measurements enable the use of combined algorithms which reduce the sensitivity to the variability of the RSD. The aim of this work is to develop a new procedure to retrieve the RSD parameters which can be used to estimate the corresponding rainfall rates. The polarimetric variables (Z_{hh}, Z_{dr}, and K_{dp}) are used to retrieve the RSD parameters by means of an ad-hoc neural network (NN) technique. The reason for this choice is the ambition to exploit the capability of NNs to approximate strongly non-linear functions such as those describing the relationships between radar observables and RSD parameters. A stochastic model, based on disdrometer measurements, is used to generate realistic range profiles of raindrop size distribution parameters while a T-matrix solution technique is adopted to compute the corresponding polarimetric variables at S band.

2 Polarimetric scattering model of rainfall

A gamma raindrop size distribution (RSD), having the general form $N(D)=N_0\, D^{\mu}\, exp(-\Lambda\, D)$ with D the particle diameter and N_0, Λ and μ RSD parameters, has been introduced in the literature to account for most of the variability occurring in the naturally observed RSD. The concept of normalization has been introduced by Willis (1984) and revisited by Chandrasekar and Bringi (1987), and Testud et al. (2001). The number of raindrops per unit volume per unit size can be written as:

$$N(D) = N_w f(\mu)\left(\frac{D}{D_0}\right)^{\mu} \exp\left[-(3.67+\mu)\frac{D}{D_0}\right] \qquad (1)$$

where $f(\mu)$ is a function μ only, the parameter D_0 is the median volume drop diameter, μ is the shape parameter of the drop spectrum, and N_w [mm^{-1} m^{-3}] is a normalized drop concentration that can be calculated as function of liquid water content W and D_0 (e.g., Bringi and Chandrasekar, 2001).

2.1 Polarimetric radar variables

The copolar radar reflectivity factors Z_{hh} and Z_{vv} [mm^6 m^{-3}] at H and V polarization state and the differential reflectivity Z_{dr} [dB] can be expressed as follows:

$$Z_{hh,vv} = \frac{\lambda^4}{\pi^5 |K|^2} < 4\pi\, |S^b_{hh,vv}(D)|^2 > \qquad (2)$$

$$Z_{dr} = 10\log_{10}\left(\frac{Z_{hh}}{Z_{vv}}\right) \qquad (3)$$

where $S_{hh,vv}$ are the backscattering co-polar components of the complex scattering matrix S of a raindrop and the angular brackets represent the ensemble average over the RSD. K depends on the complex dielectric constant of water estimated as a function of wavelength λ [mm] and temperature. For a polarimetric radar, the specific differential phase shift K_{dp} [° km^{-1}], due to the forward propagation phase difference between H and V polarization can be obtained as:

Correspondence to: Gianfranco Vulpiani.

gianfranco.vulpiani@meteo.fr

$$K_{dp} = 10^{-3}\frac{180}{\pi}\lambda \cdot \mathrm{Re}[< f_{hh}(D) - f_{vv}(D) >] \quad (4)$$

where $f_{hh,vv}$ are the forward-scattering co-polar components of S. The rainfall rate R [mm h^{-1}] is defined as

$$R = 6\pi\, 10^{-4}\int D^3\; v(D)\; N(D)\; dD \quad (5)$$

where $v(D)$ [m s^{-1}] is the raindrop terminal fall speed in still air. At sea level, a commonly used raindrop speed relationship is the one derived by Atlas and Ulbrich (1977).

3 Neural network retrieval technique

3.1 *Minimization and regularization techniques*

An artificial neural network is a non-linear parameterized mapping from an input **x** to an output **y**=*NN*(**x**; **w**, **M**), where **w** is the vector of parameters relating the input to the output while the functional form of the mapping (i.e., the architecture of the net) is denoted as **M**. The multi-layer perceptron architecture (MLP), considered here, is a mapping model composed of several layers of parallel processors. It has been theoretically proven that one-hidden layer MLP networks may represent any non-linear continuous function (Haykin, 1995), while a two-hidden layer MLP may approximate any function to any degree of non-linearity also taking into account discontinuities (Sontag, 1992). The NN architecture is such that all nodes are fully interconnected to each other and these interconnections are characterized by weights and biases. The hidden and output nodes are characterized by the activation function f which is generally assumed to be a differentiable non-linear function. Here we chose the sigmoidal function, characterized by the node gain and the node bias (Marzano et al., 2004). The network is trained using supervised learning, with a training set D=(x_i, t_i) of inputs and targets. During training, the weights (w) and biases are iteratively adjusted in order to minimize the so called network performance (objective) function, which normally is the summed squared error:

$$E_D = \frac{1}{2}\sum_{i=1}^{N}(t_i - a_i)^2 \quad (6)$$

where a_i is the neural network response. The minimization is based on repeated evaluations of the gradient of the performance function using back-propagation, which involves performing computations backwards through the network (Rumelhart, 1986). For the weights w_{jj}, that is the weight of the *i-th* output node associated to the *j-th* hidden value, the use of the delta rule leads to the following updating equation

$$\Delta w_{ij} = w_{ij}{}' - w_{ij} = -\eta_0 \frac{\partial E_D}{\partial w_{ij}} \quad (7)$$

where the w_{jj}' indicates the updated (new) value of w_{jj} and η_0 is the output-layer learning rate. The algorithm is very sensitive to the proper setting of the learning rate. For this reason, a back-propagation training with an adaptive learning rate is crucial. Battiti's "bold driver" technique (Battiti, 1989) has been implemented in this work. Gradient descent may get stuck in local minima of the performance function. The best strategy in this case is to orient the search towards the global minimum, but the form of the error function may be such that the gradient does not point in this direction. The problem can be overcome by including a momentum term m in the weight updates (Fausett, 1994). In the present study a value of m=0.9 has been assumed (Hagan, 1996). The ideal neural network is characterized by small errors on the training set and the capability to respond properly to new inputs. The latter property is called generalization. The procedure to improve generalization, called regularization, adds an additional term to the objective function which becomes $E_R = \gamma E_D + (1-\gamma)\, E_W$. E_W is the sum of squares of the networks weights and biases. In addition we note that, in Aires et al. (2002), the authors have experimentally proven that, for noisy data, a one-hidden layer MLP network may improve the network generalization through the reduction of the number of parameters. In Bishop (1996), the author demonstrated that, assuming low noise conditions, training with input perturbation is closely related to regularization.

3.2 Raindrop size distribution retrieval

Reflectivity and differential reflectivity are commonly used for RSD retrieval (Gorgucci et al., 2002; Brandes et al., 2002). Specific differential phase shift is a potential predictor for RSD estimation but it may be affected by a high noise which may perturb the results. Consequently, K_{dp} has been used in Gorgucci et al. (2002) setting a lower threshold of 0.2 deg km^{-1}. It has been found that the proposed algorithm performs well even for very low values of K_{dp}. Moreover, in case of unreliable or unavailable measurements of K_{dp} (i.e., radars which do not measure differential phase), a 2-input neural network algorithm can also be successfully used. The median volume drop diameter D_0 and the intercept parameter N_w are independently estimated using distinct NNs with 3 (i.e., Z_{hh}, Z_{dr}, K_{dp}) or 2 inputs (i.e., Z_{hh}, Z_{dr}), according to the availability and reliability of K_{dp}. The shape parameter μ is estimated from Z_{dr} and the retrieved values of D_0 (as suggested in Brandes, 2002) using 2-input (i.e., Z_{dr}, D_0) NN. Consequently, its estimate is indirectly dependent on K_{dp} through D_0. The proposed RSD retrieval technique can be formalized in the following way:

$$\begin{array}{c} NN_{D0}\left(Z_{hh}, Z_{dr}, K_{dp}\right) \\ NN_{N_W}\left(Z_{hh}, Z_{dr}, K_{dp}\right) \\ NN_{N_W}\left(Z_{dr}, \hat{D}_0\right) \end{array} \quad (8)$$

if K_{dp} is available, $NN_{D0}(Z_{hh}, Z_{dr})$ and $NN_{Nw}(Z_{hh}, Z_{dr})$ being used otherwise. The neural network architecture and regularization parameters have been determined according to a heuristic monitoring of the generalization capability on test data, the root mean square error having been used as metric. According to what is suggested in Aires (2002), it has been found that the one-hidden layer configuration improves the generalization capability of the NNs. The number of nodes in the hidden layer has been fixed to 6 for NN_{D0} and NN_{Nw}, and 12 for NN_{μ}. A choice of γ=0.7 has been found to be suitable for the retrieval of all the RSD parameters. The input perturbation technique has also been adopted in order to increase the generalization of NN_{μ} (Z_{dr}, D_0).

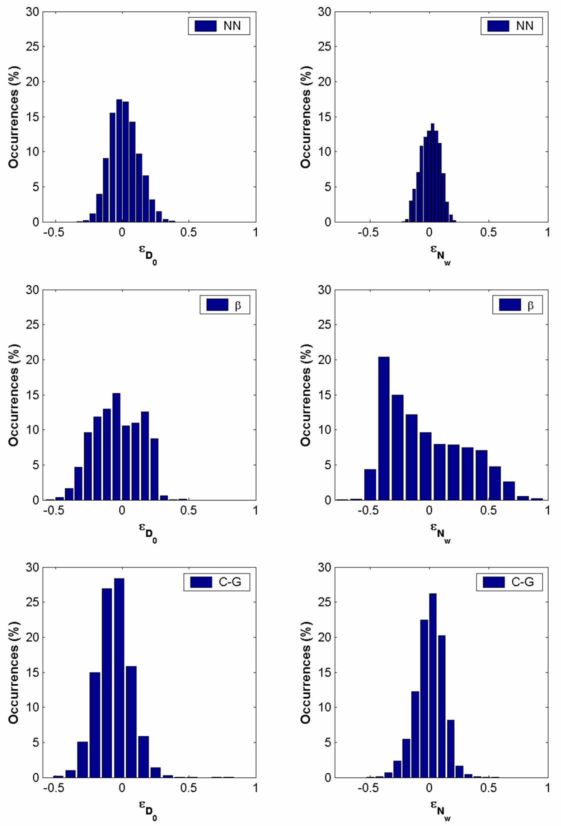

Figure 1 Error histograms of the retrieved RSD parameters. The upper panels refer to results obtained applying the NN technique. The middle and lower panels to the β and C-G methods, respectively.

4 Numerical Results on synthetic data

The choice of the architecture, learning algorithm and performance function is crucial for the neural network set up. On the other hand, the training data set also plays an important role to guaranty a high generalization capability. In order to "regularize" the neural network and test its robustness, a large highly heterogeneous data set has been generated by randomly varying the RSD parameters, the axis ratio relationship and the temperature. The validation of the proposed methodology has been accomplished using realistic range profiles of RSDs generated using a modified version of the stochastic simulator proposed by Berne and Uijlenhoet (2005). It is worth mentioning that it is based on a gamma RSD model with fixed μ (i.c., μ=3). For this reason we have focused on the retrieval of D_0 and N_w only. Once the polarimetric variables have been generated, a zero mean Gaussian noise with standard deviation equal to 1 dBZ, 0.2 dB and 0.3 ° km^{-1} has been added to Z_{hh}, Z_{dr} and K_{dp}, respectively.

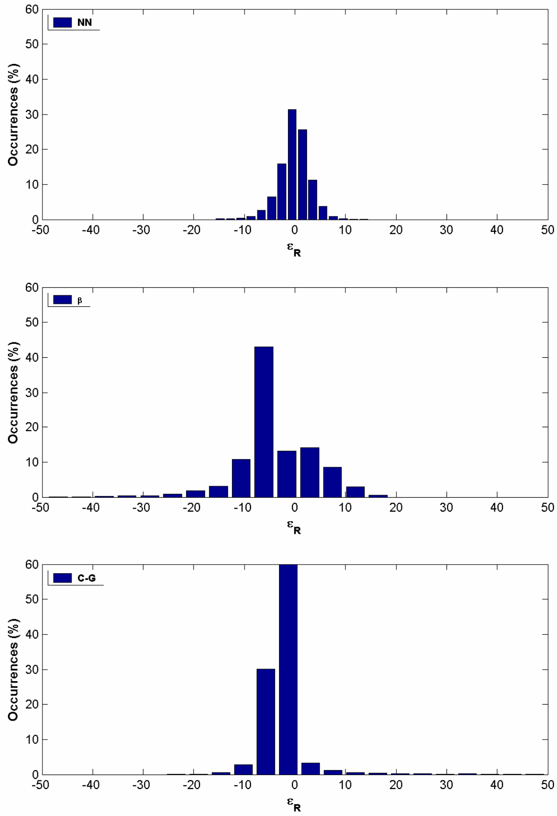

Figure 2 Error histograms of the estimated rain rate. The upper panel shows the results obtained applying the proposed NN technique. The middle and lower panels show the corresponding results for the β and C-G methods, respectively.

The sensitivity analysis has been accomplished in terms of mean error ε, root mean square error RMSE and correlation coefficient r. Figure 1 shows the comparison between the proposed technique and those described in Gorgucci et al. (2002) (named β method) and Brandes et al. (2002) (named *C-G* method). The median volume drop diameter is estimated fairly well by all examined algorithms. The differences in terms of RMSE are small but the proposed approach shows a larger correlation between the estimated and true RSD parameters. We found that $r(\hat{D}_0, D_0)$ is 0.941, 0.856 and 0.796, respectively, for the NN, C-G and β algorithms. Moreover, the estimation of N_w by means of the β algorithm is affected by a larger error standard deviation and smaller correlation coefficient as compared to the other examined methods. The latter have marked differences, especially in terms of correlation, $r(\hat{N}_w, N_w)$ being 0.943 and 0.854, respectively, for the NN and C-G methods. Once the RSD parameters are estimated using the mentioned algorithms, the corresponding rain rates can be computed using (5). R_{NN}, R_{β} and $R_{C\text{-}G}$ denote the rain rate retrieval algorithms based on the neural network, β, and C-G methods, respectively (see Figure 2). It has been found that the errors are less than 10 mm h^{-1} in about 98.2, 85.8 and 90.6% of the cases respectively for R_{NN}, R_{β}, and $R_{C\text{-}G}$. It is worth mentioning that, while the neural network based technique shows a better performance (RMSE~3.7 mm h^{-1}, r=0.951), the differences among the other algorithms are small. R_{β} provides better results in terms of RMSE (~ 8 mm h^{-1}) but a reduced correlation (r=0.758), the corresponding values for $R_{C\text{-}G}$ being about 11.4 and 0.899.

5 Conclusions

A new neural network algorithm to estimate the raindrop size distribution from S-band dual-polarized radar measurements is proposed in this work. The neural-network algorithm exhibits enhanced features to improve its efficiency, robustness and generalization capability. Numerical simulations, performed using a T-matrix hydrometeor scattering model, have been used to investigate the accuracy of the proposed methodology. The precipitation model has been characterized in terms of shape, raindrop size distribution and orientation. A disdrometer-derived stochastic model has been employed to simulate RSD variability and to construct a test data set. The error analysis, performed in order to evaluate the expected errors of this method, have shown an improvement with respect to other methodologies described in the literature to estimate RSD parameters and, consequently, the corresponding rain rates.

Acknowledgements: This work has been partially funded by European project Risk-AWARE within INTERREG-IIIB-CADSES and by the Italian Ministry of University and Research (MIUR). One of the authors (VC) acknowledges support from the NSF-ERC program (CASA-ERC-0313747). The author A.B. is financially supported by the EU Integrated Project FLOODsite (GOCE-CT-2004-505420). The author R.U. is financially supported by the Netherlands Organization for Scientific Research (NWO) through a Vernieuwingsimpuls/VIDI grant (Project 016.021.003).

References

Aires, F., A. Chédin, N. A. Scott, W. B. Rossow, 2002: A Regularized Neural Net for retrieval of Atmospheric and Surface Temperature with the IASI Instrument. *J. Appl. Meteor.*, **41**, 144-159.

Atlas D., and C. W. Ulbrich, 1977: Path and area integrated rainfall measurement by microwave attenuation in the 1-3 cm band. J. Appl. Meteor., **16**, 1322-1331.

Battiti, R., 1989: Accelerated Backpropagation Learning: Two Optimization Methods. Complex Systems, 3, 331-342.

Berne, A., and R. Uijlenhoet, 2005: A stochastic model of range profiles of raindrop size distributions: Application to radar attenuation correction. Geophys. Res. Lett., **32**, L10803, doi:10.1029/2004GL021899.

Bishop, C., 1996: Neural Networks for Pattern Recognition, Clarendon Press.

Brandes, E. A., G. Zhang, and J. Vivekanandan, 2002: Experiments in rainfall estimation with a polarimetric radar in a subtropical environment. J. Appl. Meteor., **41**, 674-685.

Chandrasekar, V., and V. N. Bringi, 1987: Simulation of radar reflectivity and surface measurements of rainfall. J. Atmos. Oceanic Technol., **4**, 464-478.

Fausett, L., 1994: Fundamentals of Neural Networks: Architectures, Algorithms and Applications. Prentice-Hall, Englewood Cliffs, NJ.

Gorgucci, E., V. Chandrasekar, V. N. Bringi, and G. Scarchilli, 2002: Estimation of raindrop size distribution parameters from polarimetric radar measurements. J. Atmos. Sci., **59**, 2373-2384.

Haykin, S., 1995: Neural Networks: A Comprehensive Foundation. Mcmillan Coll., New York, (NY).

Marzano, F. S., E. Picciotti and G. Vulpiani, 2004: Rain field and reflectivity vertical profile reconstruction from C-band radar volumetric data. IEEE Trans. Geosci. Rem. Sens., **42**, 1033-1046.

Rumelhart, D. E., G. E. Hinton, and R.J. Williams, 1986: Learning representations by back-propagating errors. Nature, **323**, 533-536.

Sontag, E., 1992: Feedback stabilization using two-hidden-layer nets. IEEE Trans. Neural Networks, **3**, 981-990.

Testud, J., S. Oury, R. A. Black, P. Amayenc, and X. Dou, 2001: The concept of “normalized” distribution to describe raindrop spectra: A tool for cloud physics and cloud remote sensing. J. Appl. Meteor., **40**, 1118-1140.

Willis, P. T., 1984: Functional fit to some observed drop size distributions and parametrization of rain. J. Atmos. Sci., **41**, 1648-1661.